

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:32:59 ; Search time 46 Seconds  
(without alignments)  
108.493 Million cell updates/sec

Title: US-10-001-938-3  
perfect score: 86  
Sequence: 1 QKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep: \*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep: \*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep: \*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep: \*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep: \*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep: \*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep: \*  
8: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep: \*  
9: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep: \*  
10: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep: \*  
11: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep: \*  
12: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep: \*  
13: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep: \*  
14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep: \*  
15: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep: \*  
16: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep: \*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep: \*  
18:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

Result No.	Score	Query %	Match Length	DB ID	Description
1	86	100.0	16	14 US-10-001-938-3	Sequence 3, Appli
2	86	100.0	340	9 US-09-764-868-1153	Sequence 1153, Appli
3	86	100.0	341	9 US-09-764-868-1152	Sequence 1152, Appli
4	86	100.0	376	15 US-10-369-493-689	Sequence 689, Appli
5	86	100.0	378	15 US-10-369-493-386	Sequence 386, Appli
6	86	100.0	378	15 US-10-369-493-21305	Sequence 21305, Appli
7	81	94.2	15	9 US-09-756-983-10	Sequence 10, Appli
8	81	94.2	15	9 US-09-828-574-10	Sequence 10, Appli
9	81	94.2	15	9 US-10-001-938-27	Sequence 27, Appli
10	81	94.2	15	14 US-10-299-540-4	Sequence 4, Appli
11	81	94.2	15	14 US-10-299-184-4	Sequence 4, Appli
12	81	94.2	15	14 US-10-239-313A-444	Sequence 444, Appli
13	81	94.2	16	14 US-10-239-313A-447	Sequence 447, Appli
14	79	91.9	365	15 US-10-369-493-10648	Sequence 10648, Appli
15	79	90.7	376	15 US-10-369-493-12087	Sequence 12087, Appli

RESULT 1  
US-10-001-938-3  
; ORGANISM: Escherichia coli

Query Match 100.0%; Score 86; DB 14; Length 16;  
Best Local Similarity 100.0%; Pred. No. 9.3e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
Db 1 QKRAAYDQYGHAAFEQ 16

**ALIGNMENTS**

RESULT 2  
US-09-764-868-1153  
; Sequence 1153, Application US/09764868  
; Patent No. US20020168711A1

; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PTZ32  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; PRIORITY APPLICATION data removed - refer to PALM or file wrapper  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1153  
; LENGTH: 340  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (110)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (135)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (156)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (196)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (198)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (216)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; US-09-764-868-1153

Query Match 100.0%; Score 86; DB 9; Length 340;  
Best Local Similarity 100.0%; Pred. No. 2.2e-06; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OKRAAYDQYGHAAFEQ 16  
Db 25 OKRAAYDQYGHAAFEQ 40

RESULT 3  
US-09-764-868-1152  
Sequence 1152, Application US/09764868  
Patent No. US20020168711A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PTZ32  
CURRENT APPLICATION NUMBER: US/09/764,868  
PRIORITY APPLICATION NUMBER: US/09/764,868  
PRIORITY FILING DATE: 2001-01-17  
NUMBER OF SEQ ID NOS: 1510  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1152  
LENGTH: 341  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-868-1152

Query Match 100.0%; Score 86; DB 9; Length 341;  
Best Local Similarity 100.0%; Pred. No. 2.2e-06; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OKRAAYDQYGHAAFEQ 16  
Db 26 OKRAAYDQYGHAAFEQ 41

RESULT 4  
US-10-369-493-689

Query Match 100.0%; Score 86; DB 9; Length 340;  
Best Local Similarity 100.0%; Pred. No. 2.2e-06; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OKRAAYDQYGHAAFEQ 16  
Db 61 OKRAAYDQYGHAAFEQ 76

RESULT 5  
US-10-369-493-386  
Sequence 386, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURREN T FILING DATE: 2003-02-28  
PRIORITY APPLICATION NUMBER: US 60/360,039  
PRIORITY FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 386  
LENGTH: 378  
TYPE: PRT  
ORGANISM: Xenorhabdus nematophilus  
US-10-369-493-386

Query Match 100.0%; Score 86; DB 15; Length 378;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OKRAAYDQYGHAAFEQ 16  
Db 61 OKRAAYDQYGHAAFEQ 76

RESULT 6  
US-10-369-493-21305  
Sequence 21305, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: July 11, 2004, 09:30:33 ; Search time 23 Seconds  
 (without alignments)  
 35.914 Million cell updates/sec

Title: US-10-001-938-3  
 Perfect score: 86  
 Sequence: 1 QKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Issued\_Patents\_AA:  
 1: /cggn2\_6/ptodata/2/iaa/5A\_COMB.pep:  
 2: /cggn2\_6/ptodata/2/iaa/5B\_COMB.pep:  
 3: /cggn2\_6/ptodata/2/iaa/6A\_COMB.pep:  
 4: /cggn2\_6/ptodata/2/iaa/6B\_COMB.pep:  
 5: /cggn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:  
 6: /cggn2\_6/ptodata/2/iaa/backfilesl.pep:  
 ;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	86	100.0	131	4	US-09-553-498-4
2	86	100.0	131	4	US-09-618-869-4
3	86	100.0	380	4	US-09-543-681A-5912
4	86	100.0	380	4	US-09-489-039A-9011
5	86	100.0	399	4	US-09-553-498-2
6	86	100.0	399	4	US-09-618-869-2
7	81	94.2	15	1	US-08-618-464-4
8	81	94.2	15	3	US-09-107-615-4
9	81	94.2	15	5	PCT-US95-04896-4
10	72	83.7	15	1	US-08-618-464-5
11	72	83.7	15	3	US-09-107-615-5
12	72	83.7	15	5	PCT-US95-04896-5
13	62	72.1	381	4	US-09-252-991A-27174
14	61	70.9	407	4	US-09-540-236-2899
15	56	65.1	352	2	US-08-472-534-6
16	54	62.8	15	1	US-08-618-464-6
17	54	62.8	15	3	US-09-107-615-6
18	52	60.5	375	4	US-09-328-352-4984
19	47	54.7	385	4	US-09-134-001C-3688
20	46	53.5	419	2	US-08-686-417-3
21	45	52.3	677	4	US-09-252-991A-18102
22	44	51.2	915	4	US-09-252-991A-24992
23	43	50.0	1895	2	US-08-619-554-4
24	41	47.7	87	4	US-09-882-835-4
25	41	47.7	223	4	US-09-658-644-4
26	41	47.7	348	2	US-08-974-546-1
27	41	47.7	4	US-09-543-681A-7130	

## ALIGNMENTS

RESULT 1  
 US-09-553-498-4  
 ; Sequence 4, Application US/09553498  
 ; Patent No. 6309861

## GENERAL INFORMATION:

; APPLICANT: Ambrosius, Dorothée  
 ; APPLICANT: Rudolph, Rainer  
 ; APPLICANT: Schaeffner, Joerg  
 ; APPLICANT: Schwarz, Elisabeth  
 ; TITLE OF INVENTION: Process for the production of naturally folded and secreted proteins  
 ; FILE REFERENCE: Case 20379  
 ; CURRENT APPLICATION NUMBER: US/09/553, 498  
 ; CURRENT FILING DATE: 2000-04-20  
 ; PRIOR APPLICATION NUMBER: EP99107412.1  
 ; PRIOR FILING DATE: 1999-04-26  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SEQ ID NO 4  
 ; LENGTH: 131  
 ; TYPE: PRT  
 ; ORGANISM: E. coli

US-09-553-498-4  
 ; Query Match %  
 ; Best Local Similarity 100.0%; Score 86; DB 4; Length 131;  
 ; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
 Db 84 QKRAAYDQYGHAAFEQ 99

RESULT 2  
 US-09-618-869-4  
 ; Sequence 4, Application US/09618869  
 ; Patent No. 6455279

; GENERAL INFORMATION:  
 ; APPLICANT: Ambrosius, Dorothée  
 ; APPLICANT: Rudolph, Rainer  
 ; APPLICANT: Schaeffner, Joerg  
 ; APPLICANT: Schwarz, Elisabeth  
 ; TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR CHAPERONES  
 ; FILE REFERENCE: 20381

; CURRENT APPLICATION NUMBER: US/09/618, 869

; CURRENT FILING DATE: 2000-07-19

; TITLE OF INVENTION: CHAPERONES

; PRIOR FILING DATE: 1999-07-29

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: Patentin Ver. 2.1

RESULT 5  
; SEQ ID NO 4  
; LENGTH: 131  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-618-869-4

Query Match 100.0%; Score 86; DB 4; Length 131;  
Best Local Similarity 100.0%; Pred. No. 7.8e-08; Mismatches 0; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
Db 84 QKRAAYDQYGHAAFEQ 99

RESULT 3  
US-09-543-681A-5912  
; Sequence 5912, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: Case 20379  
; CURRENT APPLICATION NUMBER: US/09/543,681A.  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 5912  
; LENGTH: 380  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-5912

Query Match 100.0%; Score 86; DB 4; Length 380;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07; Mismatches 0; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
Db 63 QKRAAYDQYGHAAFEQ 78

RESULT 4  
US-09-489-039A-9011  
; Sequence 9011, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.200401  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 9011  
; LENGTH: 380  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9011

Query Match 100.0%; Score 86; DB 4; Length 380;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07; Mismatches 0; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
Db 64 QKRAAYDQYGHAAFEQ 79

RESULT 6  
US-09-618-869-2  
; Sequence 2, Application US/09618869  
; Patent No. 6455279  
; GENERAL INFORMATION:  
; APPLICANT: Ambrosius, Dorthee  
; APPLICANT: Rudolph, Rainer  
; APPLICANT: Schaeffner, Joerg  
; APPLICANT: Schwarz, Elisabeth  
; TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR CHAPERONES  
; FILE REFERENCE: 20381  
; CURRENT APPLICATION NUMBER: US/09/618,869  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: EP99114811.5  
; PRIOR FILING DATE: 1999-07-29  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-618-869-2

Query Match 100.0%; Score 86; DB 4; Length 399;  
Best Local Similarity 100.0%; Pred. No. 2.5e-07; Mismatches 0; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
Db 84 QKRAAYDQYGHAAFEQ 99

RESULT 7  
US-08-618-464-4  
; Sequence 4, Application US/08618464  
; Patent No. 573570  
; GENERAL INFORMATION:  
; APPLICANT: CARSON, DENNIS A.  
; APPLICANT: ALBANI, SALVATORE  
; TITLE OF INVENTION: VACCINE COMPOSITIONS AND METHODS USEFUL IN

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:25:58 ; Search time 11 seconds  
 (without alignments)  
 75.738 Million cell updates/sec

Title: US-10-001-938-3  
 Perfect score: 86  
 Sequence: 1 QKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SwissProt\_42;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	86	100.0	375	1 DNAJ_ECOLI
2	86	100.0	377	1 DNAJ_HAEDU
3	86	100.0	378	1 DNAJ_SALTY
4	86	100.0	382	1 DNAJ_HAEIN
5	83	96.5	375	1 DNAJ_ACTAC
6	83	96.5	381	1 DNAJ_VIBCH
7	83	96.5	385	1 DNAJ_VIBHA
8	82	95.3	392	1 NOLC_RHIFR
9	80	93.0	372	1 DNAJ_PASMU
10	77	89.5	234	1 DNAJ_RHILE
11	77	89.5	379	1 DNAJ_RHOSP
12	75	87.2	377	1 DNAJ_BUCAI
13	75	87.2	379	1 DNAJ_PASHA
14	75	87.2	383	1 DNAJ_BUCBP
15	74	86.0	378	1 DNAJ_BUCAP
16	73	84.9	370	1 DNAJ_ERYRH
17	73	84.9	373	1 DNAJ_NEIMA
18	73	84.9	375	1 DNAJ_BRUOV
19	73	84.9	377	1 DNAJ_AGRTS
20	73	84.9	377	1 DNAJ_BRAJA
21	73	84.9	377	1 DNAJ_BRUME
22	73	84.9	377	1 DNAJ_BRUSU
23	69	80.2	384	1 DNAJ_RHOCA
24	67	77.9	371	1 DNAJ_METSS
25	66	76.7	383	1 DNAJ_LACSK
26	64	74.4	370	1 DNAJ_RICPR
27	64	74.4	374	1 DNAJ_COXBU
28	63	73.3	1 DNJ2_AQURE	
29	63	73.3	379	1 DNAJ_LEGPN
30	62	72.1	377	1 DNAJ_PSEAE
31	58	67.4	379	1 DNAJ_LACLA
32	58	67.4	385	1 DNAJ_CAUCR
33	57	66.3	1 DNAJ_BORBU	

006431 nitrosomona  
 Q92bng listeria in  
 Q9s5a3 listeria mo  
 Q9pb06 xylella fas  
 Q87bs9 xylella fas  
 P95830 streptococc  
 Q9xca6 porphyromon  
 O34136 deinococcus  
 P48207 francisella  
 P17631 bacillus su  
 Q27352 methanobact  
 Q9hhb8 halobacteri

#### ALIGNMENTS

RESULT 1	ID	NAME	STANDARD	PRT;	375 AA.
	ID	DNAJ_ECOLI			
	AC	P08622;			
	DT	01-AUG-1988 (Rel. 08, Created)			
	DT	15-DEC-1998 (Rel. 37, Last sequence update)			
	DT	28-FEB-2003 (Rel. 41, Last annotation update)			
	DE	Chaperone protein dnaJ (Heat shock protein J) (HSP40).			
	GN	DNAJ OR GROP OR B0015.			
	OS	Escherichia coli.			
	OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
	OC	Enterobacteriaceae; Escherichia.			
	OX	NCBI_TaxID=562;			
	RN	[1]			
	RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-5.			
	RX	MEDLINE=86111849; PubMed=3003084;			
	RA	Onki M., Tamura F., Nishimura S., Uchida H.;			
	RT	"Nucleotide sequence of the Escherichia coli dnaJ gene and purification of the gene product"; J. Biol. Chem. 261:1778-1781(1986).			
	RL				
	RN	[2]			
	RP	SEQUENCE FROM N.A.			
	RC	SEQUENCE FROM N.A.			
	RX	MEDLINE=86111850; PubMed=3003085;			
	RA	Bardwell J.C.A., Tilly K., Craig E., King J., Zylizc M., Georgopoulos C.;			
	RA	"The nucleotide sequence of the Escherichia coli K12 dnaJ+ gene. A gene that encodes a heat shock protein. "; J. Biol. Chem. 261:1782-1785(1986).			
	RT				
	RL				
	RN	[3]			
	RP	SEQUENCE FROM N.A.			
	RC	SEQUENCE FROM N.A.			
	RX	MEDLINE=92334977; PubMed=1630901;			
	RA	Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N., Isono K., Mizobuchi K., Nakata A.;			
	RA	"Systematic sequencing of the Escherichia coli genome: analysis of the 0-2.4 min region. "; Nucleic Acids Res. 20:3305-3308(1992).			
	RT				
	RL				
	RN	[4]			
	RP	SEQUENCE FROM N.A.			
	RC	SEQUENCE FROM N.A.			
	RX	MEDLINE=97426617; PubMed=9278503;			
	RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;			
	RA	"The complete genome sequence of Escherichia coli K-12. "; Science 277:1453-1474 (1997).			
	RT				
	RL				
	RN	[5]			
	RP	POSSIBLE FUNCTION.			
	RX	MEDLINE=9118784; PubMed=1826368;			
	RA	Liberek K., Marszalek J., Ang D., Georgopoulos C., Zylizc M.; "Escherichia coli DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. "; Proc. Natl. Acad. Sci. U.S.A. 88:2874-2878(1991).			
	RL	[6]			



Copyright (c) 1993 - 2004 Compugen Ltd.  
 OM protein - protein search, using sw model  
 Run on: July 11, 2004, 09:29:38 ; Search time 20 seconds  
 Perfect score: 86 US-10-001-938-3  
 Sequence: 1 QKRAAYDQYGHAAFEQ 16  
 Scoring table: BLOSUM62 GapOp 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : PIR\_78:\*

- 1: pir1:\*
- 2: pir2:\*
- 3: pir3:\*
- 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No. Score Query Match Length DB ID

Result No.	Score	Query	Match	Length	DB	ID	Description
1	86	100.0	376	1	HHECDJ		heat shock protein dnaJ - Escherichia coli (strain K-12)
2	86	100.0	376	2	G90630		heat shock protein dnaJ protein [impo
3	86	100.0	376	2	G85481		hypothetical protein
4	86	100.0	379	2	AB0058		heat shock protein
5	86	100.0	379	2	AF0503		heat shock protein
6	86	100.0	394	2	C64112		probable curved-DN
7	83	96.5	381	2	D82270		heat shock protein
8	82	95.3	392	2	S15295		DnaJ protein [impo
9	75	87.2	377	2	JC5609		heat shock protein
10	75	87.2	377	2	F84947		chaperone with Dna
11	73	84.9	373	2	D81242		chaperone protein
12	73	84.9	375	2	B47042		DnaJ protein [impo
13	73	84.9	377	2	AC3502		heat shock protein
14	73	84.9	377	2	AD2591		DnaJ protein NMA02
15	73	84.9	377	2	D97373		heat shock protein
16	64	74.4	367	2	I140843		heat shock protein
17	64	74.4	370	2	C71729		DnaJ protein - Rhi
18	63	73.3	376	2	E70361		heat shock protein
19	62	72.1	190	2	B35388		chaperone protein
20	62	72.1	377	2	A83076		molecular chaperon
21	60	69.8	373	2	H97728		chaperone protein
22	58	67.4	379	2	A47079		heat shock protein
23	58	67.4	379	2	H86902		DnaJ protein [impo
24	58	67.4	385	2	C67250		heat shock protein
25	57	66.3	352	2	A49210		DnaJ protein [impo
26	57	66.3	364	2	D70164		heat shock protein
27	57	66.3	376	2	AD1621		heat shock protein
28	57	66.3	377	2	T43739		F;144-151/Region: CXCXGXG repeat
29	57	66.3	377	2	AH1258		F;161-168/Region: CXCXGXG repeat

#### ALIGNMENTS

30	56	65.1	368	2	F82570	DnaJ protein XF233
31	56	65.1	372	2	H97928	heat-shock protein
32	56	65.1	378	2	D95060	dnaJ protein [impo
33	53	61.6	499	2	G96831	hypothetical prote
34	52	60.5	372	2	B41874	heat shock protein
35	52	60.5	376	2	H69038	heat shock protein
36	52	60.5	389	2	S41748	heat shock protein
37	50	58.1	297	2	G81329	probable curved-DN
38	50	58.1	348	2	T04618	heat shock protein
39	50	58.1	379	2	B89939	DnaJ protein [impo
40	50	58.1	383	2	E90603	heat shock protein
41	50	58.1	416	2	F71379	heat shock protein
42	49	57.0	335	2	T48161	heat shock protein
43	49	57.0	346	2	B84602	probable DnaJ prot
44	49	57.0	370	2	D83818	heat-shock protein
45	49	57.0	375	2	D82894	heat shock protein

F;183-190/Region: CXXCXGXG repeat  
F;197-204/Region: CXXCXGXG repeat

Query Match Best Local Similarity 100.0%; Score 86; DB 1; Length 376;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
Db 61 QKRAAYDQYGHAAFEQ 76

RESULT 2  
G90630  
DnaJ protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C;Accession: G90630  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuwara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: G90630  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-376 <HAY>  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
A;Gene: ECs0015  
C;Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology  
Query Match Best Local Similarity 100.0%; Score 86; DB 2; Length 376;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
Db 61 QKRAAYDQYGHAAFEQ 76

RESULT 3  
G85481  
Chaperone with DnaK, heat shock protein [imported] - Escherichia coli (strain O157:H7, s  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C;Accession: G85481  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimaranta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: G85481  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-376 <STO>  
A;Cross-references: GB:AE005174; NID:912512693; PIDN:AAG54315.1; GSPDB:GN00145; UWGP:Z00  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: dnaJ  
C;Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology  
Query Match Best Local Similarity 100.0%; Score 86; DB 2; Length 376;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
Db 61 QKRAAYDQYGHAAFEQ 76

RESULT 4  
C;Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology  
Query Match Best Local Similarity 100.0%; Score 86; DB 2; Length 376;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
Db 61 QKRAAYDQYGHAAFEQ 76

RESULT 4

RESULT 5  
AF0503  
DnaJ protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18)  
C;Species: Salmonella enterica subsp. enterica serovar Typhi  
A;Note: this species has also been called Salmonella typhi  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C;Accession: AF0503  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
Th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
'S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A;Title: Complete genome sequence of a multiple drug resistant Salmonella serovar  
A;Reference number: AB0502; MUID:21534947; PMID:11677608  
A;Accession: AF0503  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-379 <PAR>  
A;Cross-references: GB:AL513382; PIDN:CAD01166.1; PID:916501296; GSPDB:GN00176  
C;Genetics:  
A;Gene: STY0013  
C;Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology  
Query Match Best Local Similarity 100.0%; Score 86; DB 2; Length 379;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
Db 61 QKRAAYDQYGHAAFEQ 76

RESULT 6  
C64112  
heat shock protein dnaJ - Haemophilus influenzae (strain Rd KW20)  
C;Species: Haemophilus influenzae  
C;Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 20-Aug-1999  
C;Accession: C64112  
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.  
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: July 11, 2004, 10:05:13 ; Search time 54 Seconds  
 (without alignments)  
 83.718 Million cell updates/sec

Title: US-10-001-938-3

Perfect score: 86

Sequence: 1 QKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp19805:\*

2: geneseqp19905:\*

3: geneseqp20005:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	86	100.0	16 5 AAU98852	Aau98852 E.Coli DN
2	86	100.0	16 6 ABR55126	Abr55126 E. coli d
3	86	100.0	131 3 AAB11396	Aab11396 E. coli e
4	86	100.0	131 4 AAB74197	Aab74197 OmpA-J do
5	86	100.0	131 4 AAB70767	Aab70767 Expressio
6	86	100.0	131 4 AAY72018	Aay72018 E. coli O
7	86	100.0	340 4 AAU17588	Aau17588 Novel sig
8	86	100.0	340 7 ADB94296	Adb94296 Human nov
9	86	100.0	341 4 AAU17587	Aau17587 Novel sig
10	86	100.0	341 7 ADB94295	Adb94295 Human nov
11	86	100.0	373 6 ABM67485	Abm67485 Photorhab
12	86	100.0	399 3 AAB11395	Aab11395 E. coli e
13	86	100.0	399 4 AAB74196	Aab74196 OmpA-DnaJ
14	86	100.0	399 4 AAB70766	Aab70766 Expressio
15	86	100.0	399 4 AAY72017	Aay72017 E. coli O
16	86	100.0	476 4 ABG1771	Abg1771 Novel hum
17	81	94.2	15 2 AAW25795	AAW25795 Peptide d
18	81	94.2	15 3 AAY79543	Aay79543 E. coli d
19	81	94.2	15 4 AAM9341	Aam9341 Vaccine r
20	81	94.2	15 5 AAU09842	AAU09842 E. coli dn
21	81	94.2	15 5 AAU98876	Aau98876 E. coli D
22	81	94.2	15 5 AAE19457	Aae19457 Human hea
23	81	94.2	15 6 ABR55132	Abr55132 E. coli D
24	81	94.2	16 4 AAM9344	Aam9344 Vaccine r
25	81	94.2	18 4 AAU02075	Aau02075 Synthetic

This invention relates to the use of a peptide, which is an immunogenic portion derived from a dnaJ heat shock protein (hsp) in modulating an immune response in a subject. The peptides of the invention may have immunomodulatory, cytostatic, antibacterial or antiarthritic properties and can stimulate expression of interleukins, tumour necrosis factor and transforming growth factor beta. The immunogenic peptide is useful for modulating (i.e. augmenting/inducing or reducing/inhibiting) an immune response in a subject having an

CC immunological disorder (e.g. autoimmune disease such as arthritis or CC articular juvenile idiopathic arthritis), an infectious disease, an CC inflammatory bowel disease or cancer. The immunogenic peptide of the CC invention is also useful for modulating immunoefector cell CC responsiveness in a subject. The immunogenic peptide is particularly useful for treating the above-mentioned diseases in mammals, e.g. cat, CC dog, horse, farm animal (e.g. ovine, bovine or porcine) or human. In general, the peptide is useful in methods involving mucosal tolerisation, CC DNA vaccination, energy induction or active immunisation. The present sequence represents an E. coli DNA<sub>J</sub> immunogenic peptide of the invention CC SQ Sequence 16 AA;

Query Match 100.0%; Score 86; DB 5; Length 16;  
Best Local Similarity 100.0%; Pred. No. 4.1e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QKRAAYDQYGHAAFEQ 16  
Db 1 QKRAAYDQYGHAAFEQ 16

RESULT 2  
ABR55126  
ID ABR55126 standard; peptide; 16 AA.

XX AC ABR55126;  
XX DT 03-JUL-2003 (first entry)  
XX DE E. coli dnaJ61 antigen-specific epitope peptide.  
XX KW Antigen-specific epitope; immune response; T cell; cytokine;  
KW antiarthritic; antidiabetic; neuroprotective; anti-inflammatory;  
KW cytostatic; antithyroid; anti-asthmatic; immunosuppressive;  
KW antipsoriatic; anti-ulcer; antianæmic; cardiant; respiratory;  
KW antiallergic; dermatological; antipsoriatic.

XX OS Escherichia coli.  
XX OS Synthetic.  
XX PN EP1048732-A1.  
XX PD 02-NOV-2000.  
XX PF 26-APR-1999; 99EP-00107412.  
XX PR 26-APR-1999; 99EP-00107412.  
XX PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.  
XX DR WO2003026579-A2.  
XX PN 03-APR-2003.  
XX PR 25-SEP-2002; 2002WO-US030578.  
XX PR 25-SEP-2001; 2001US-0325499P.  
XX PR 11-DEC-2001; 2001US-0339284P.  
XX PA (REGC ) UNIV CALIFORNIA.  
XX PI Albani S, Martins A;  
XX DR WPI; 2000-674185/66.  
XX DR N-PSDB; AAC66072.

XX PT Preparation of water-soluble eukaryotic polypeptides with disulfide PT bridges e.g. rPA, comprises cultivation of prokaryotic cells in the PT presence of arginine or amide compound.

XX PS Example 3; Page 17; 40pp; German.

CC This invention describes a novel preparation of a water-soluble CC naturally occurring eukaryotic polypeptide containing two or more CC cysteine units bound via a disulfide bridge which comprises cultivation CC of prokaryotic cells in the presence of arginine or an amide compound. CC The method is useful for the preparation of eukaryotic proteins e.g. CC proteases, interferons, protein hormones, antibodies or antibody CC fragments (e.g. a single chain Fv fragment that binds to thyroid CC stimulating hormone). It is especially useful for preparing proteins with CC more than five disulfide bridges, e.g. recombinant plasminogen activator CC (rPA). The technique is simple and does not require in vitro after- CC treatment, such as the removal of inclusion bodies, reduction or CC naturalization CC SQ Sequence 131 AA;

Query Match 100.0%; Score 86; DB 3; Length 131;  
Best Local Similarity 100.0%; Pred. No. 4e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QKRAAYDQYGHAAFEQ 16  
Db 84 QKRAAYDQYGHAAFEQ 99

RESULT 4

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:41:15 ; Search time 44 seconds

(without alignments)  
 113.425 Million cell updates/sec

Title: US-10-001-938-3  
 Perfect score: 86  
 Sequence: 1 OKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 218390

Minimum DB seq length: 0  
 Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : published\_Applications\_AA:\*

```

1: /cggn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cggn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cggn2_6/ptodata/1/pubpaa/US06_NEW_PUBCOMB.pep:*
4: /cggn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5: /cggn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
6: /cggn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cggn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cggn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cggn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cggn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cggn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cggn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cggn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cggn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cggn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cggn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cggn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cggn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	86	100.0	16	14	US-10-001-938-3
2	81	94.2	15	9	US-09-756-983-10
3	81	94.2	15	9	US-09-828-574-10
4	81	94.2	15	14	US-10-001-938-27
5	81	94.2	15	14	US-10-299-540-4
6	81	94.2	15	14	US-10-299-184-4
7	81	94.2	15	14	US-10-239-313A-444
8	81	94.2	15	14	US-10-239-313A-447
9	72	83.7	15	9	US-09-828-574-11
10	72	83.7	15	14	US-10-299-540-5
11	72	83.7	15	14	US-10-299-184-5
12	54	62.8	15	14	US-10-299-540-6
13	54	62.8	15	14	US-10-299-184-6
14	39.5	9	14	US-10-283-423-154	
15	39.5	9	14	US-10-283-423-155	

RESULT 1  
 US-10-001-938-3  
 Sequence 3, Application US/10001938  
 Publication No. US20030031679A1

GENERAL INFORMATION:

APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
 APPLICANT: ALBANI, Salvatore  
 APPLICANT: CARSON, Dennis  
 APPLICANT: PRAKERN, Berent  
 APPLICANT: MARTINI, Alberto  
 TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES DERIVED FROM HEAT SHOCK PROTEINS AND USI  
 TITLE OF INVENTION: THEREOF  
 FILE REFERENCE: UCSD1360-1

CURRENT APPLICATION NUMBER: US/10/001,938

CURRENT FILING DATE: 2001-10-31

PRIOR APPLICATION NUMBER: US 60/245,181

PRIOR FILING DATE: 2000-11-01

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn version 3.0

; SEQ ID NO 3  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Escherichia coli

US-10-001-938-3

Query Match 100.0%; Score 86; DB 14; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 9.3e-08;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1 OKRAAYDQYGHAAFEQ 16	1 OKRAAYDQYGHAAFEQ 16
Db	1 OKRAAYDQYGHAAFEQ 16	1 OKRAAYDQYGHAAFEQ 16

RESULT 2  
 US-09-756-983-10  
 Sequence 10, Application US/09756983  
 Patent No. US20020122818A1

```

; GENERAL INFORMATION:
; APPLICANT: Albani, Salvatore
; TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,
; TITLE OF INVENTION: CHARACTERIZATION AND MODULATION OF
; TITLE OF INVENTION: ANTIGEN-SPECIFIC T CELLS
; CURRENT APPLICATION NUMBER: US/09/756,983
; CURRENT FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/105,018
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 09/421,506
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: PCT/US99/2466
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 15
; TYPE: PRT
; ORGANISM: E. coli
; FEATURE:
; OTHER INFORMATION: dnaJpl heat shock protein
; US-09-756-983-10

Query Match          94.2%; Score 81; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QKRAAYDQYGHAAFE 15
Db      1 QKRAAYDQYGHAAFE 15

RESULT 3
US-09-828-574-10
; Sequence 10, Application US/09828574
; Patent No. US20020146759A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: ALBANI, Salvatore
; APPLICANT: PRAKKEN, Berent J.
; TITLE OF INVENTION: STRESS PROTEINS AND PEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: UCSD1310-1
; CURRENT APPLICATION NUMBER: US/09/828,574
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/224,104
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: dnaJPI peptide
; US-09-828-574-10

Query Match          94.2%; Score 81; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QKRAAYDQYGHAAFE 15
Db      1 QKRAAYDQYGHAAFE 15

RESULT 4
US-10-001-938-27
; Sequence 27, Application US/10001938
; Publication No. US20030031679A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: ALBANI, Salvatore

RESULT 5
US-10-299-540-4
; Sequence 4, Application US/10299540
; Publication No. US20030143238A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: CARSON, Dennis A.
; APPLICANT: ALBANI, Salvatore
; TITLE OF INVENTION: VACCINE COMPOSITIONS AND METHODS USEFUL IN INDUCING IMMUNE PROTECTION AGAINST ARTHRITIS
; TITLE OF INVENTION: ARTHRITIS
; FILE REFERENCE: UCSD1370-7
; CURRENT APPLICATION NUMBER: US/10/299,540
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US 09/616,247
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/107,615
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 08/618,464
; PRIOR FILING DATE: 1996-03-15
; PRIOR APPLICATION NUMBER: US 08/246,988
; PRIOR FILING DATE: 1994-05-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic dnaJpl
; US-10-299-540-4

Query Match          94.2%; Score 81; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QKRAAYDQYGHAAFE 15
Db      1 QKRAAYDQYGHAAFE 15

RESULT 6
US-10-299-184-4
; Sequence 4, Application US/10299184
; Publication No. US20030147910A1
; GENERAL INFORMATION:

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:38:55 ; Search time 20 seconds  
(without alignments)  
76.953 Million cell updates/sec

Title: US-10-001-938-3  
Perfect score: 86  
Sequence: 1 OKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 2772

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing First 45 summaries

Database : PIR\_78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

Result No.	Score	Query Match Length	DB ID	Description
1	34	39.5	11	2 A60656 perisulfakinin - A
2	34	39.5	14	2 A56632 neosulfakinin-II -
3	31	36.0	10	1 GMROL2 leucosulfakinin-II
4	31	36.0	10	2 B60656 leucosulfakinin II
5	31	36.0	13	2 S47372 T-cell antigen rec
6	31	36.0	13	2 S47384 T-cell antigen rec
7	30	34.9	11	1 GMROL leucosulfakinin-II
8	30	34.9	13	2 S47368 T-cell antigen rec
9	28	32.6	13	2 S47359 T-cell antigen rec
10	28	32.6	13	2 S47365 T-cell antigen rec
11	28	32.6	13	2 S47374 T-cell receptor be
12	26	30.2	11	2 PT0214 T-cell receptor be
13	26	30.2	13	2 S47388 T-cell receptor be
14	26	30.2	13	2 PH1479 T-cell receptor be
15	26	30.2	14	4 I52618 hemoglobin beta ch
16	26	30.2	15	2 PH0770 T-cell receptor be
17	25	29.1	13	2 S32471 lymnaDFamide 1 - g
18	25	29.1	14	2 PH1586 Ig H chain V-D-J r
19	24	27.9	16	2 PH1622 Ig H chain V-D-J r
20	24	27.9	16	2 S11078 glucose-6-phosphate
21	23	26.7	13	2 S32472 lymnaDFamide 2 - g
22	23	26.7	13	2 A60379 factor X activator
23	26	26.7	13	2 S32473 lymnaDFamide 3 - g
24	22	25.6	13	2 S47383 T-cell antigen rec
25	22	25.6	15	2 PH1319 Ig heavy chain DJ
26	22	25.6	15	2 F28587 T-cell receptor be
27	22	25.6	15	2 I53284 T-cell receptor be
28	22	25.6	15	2 PT0215 T-cell receptor be

**ALIGNMENTS**

Query	Match	Best Local Similarity	Score	DB	Length
QY	6 YDQYGHAAF	39.5%	34	2	11
Db	3 FDDYGHMRF	55.6%	9	1	11

**RESULT 1**  
A60656 perisulfakinin - American cockroach  
C;Species: Periplaneta americana (American cockroach)  
C;Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 11-Jul-1997  
C;Accession: A60656  
R;Veenstra, J.A.  
Neuropeptides 14, 145-149, 1989  
A;Title: Isolation and structure of two gastrin/CCK-like neuropeptides from the American  
A;Reference number: A60656; MUID:90137190; PMID:2615921  
A;Accession: A60656  
A;Molecule type: protein  
A;Residues: 1-11 <VEE>  
C;Comment: This neuropeptide stimulates hindgut contractions.  
C;Keywords: amidated carboxyl end; neuropeptide; sulfoprotein  
F;6/Binding site: sulfate (Tyr) (covalent) #status experimental  
F;11/Modified site: amidated carboxyl end (Phe) #status experimental

**RESULT 2**  
A56632 neosulfakinin-II - flesh fly (*Sarcophaga bullata*)  
N;Alternate names: Neb-SK-II  
N;Contains: neosulfakinin-I (Neb-SK-I)  
C;Species: *Sarcophaga bullata*  
C;Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 20-Jun-2000  
C;Accession: A56632  
C;Fonagy, A.; Schoofs, L.; Proost, P.; Van Damme, J.; De Loof, A.  
Comp. Biochm. Physiol. C 103, 135-142, 1992  
A;Title: Isolation and primary structure of two sulfakinin-like peptides from the fleshfly  
A;Reference number: A56632; MUID:93083101; PMID:1360367  
A;Accession: A56632  
A;Molecule type: protein  
A;Residues: 1-14 <FON>  
A;Experimental source: heads  
A;Note: sequence extracted from NCBI backbone (NCBIP:120391)  
C;Keywords: amidated carboxyl end; neuropeptide; sulfoprotein  
F;1-14/Product: neosulfakinin-II #status experimental <NSK2>  
F;6-14/Product: neosulfakinin-I #status experimental <NSK1>  
F;9/Binding site: sulfate (Tyr) (covalent) #status predicted  
F;14/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match Best Local Similarity 39.5%; Score 34; DB 2; Length 14;  
Best Local Similarity 55.6%; Pred. No. 12;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Qy 6 YDQYGHAAF 14  
 C;Species: Homo sapiens (man)  
 C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999  
 C;Accession: S47372  
 R;Nachman, R.J.; Holman, G.M.; Cook, B.J.; Haddon, W.F.; Ling, N.  
 Biochem. Biophys. Res. Commun. 140, 357-364, 1986  
 A;Title: Leucosulfakinin-II, a blocked sulfated insect neuropeptide with homology to chd  
 A;Reference number: A26335; MUID:87048769; PMID:3778455  
 A;Molecule type: protein  
 A;Residues: 1-10 <NAC>  
 C;Comment: This peptide was isolated from head extracts. It stimulates muscle contractic  
 C;Superfamily: gastrin  
 C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid; sulfoprotein  
 F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F;5/Binding site: sulfate (Tyr) (covalent) #status experimental  
 F;10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 36.0%; Score 31; DB 1; Length 10;  
 Best Local Similarity 62.5%; Pred. No. 27;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 7 DQYGHAAF 14  
 Db 3 DDYGHMRF 10

---

RESULT 4

B60656  
 Leucosulfakinin II, non-sulfated - American cockroach  
 C;Species: Periplaneta americana (American cockroach)  
 C;Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 08-Dec-1995  
 C;Accession: B60656  
 R;Veenstra, J.A.  
 Neuropeptides 14, 145-149, 1989  
 A;Title: Isolation and structure of two gastrin/CCK-like neuropeptides from the American  
 A;Reference number: A60656; MUID:90137190; PMID:2615921  
 A;Accession: B60656  
 A;Molecule type: protein  
 A;Residues: 1-10 <VEE>  
 C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid (Gln) #status experimental  
 F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F;10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 36.0%; Score 31; DB 2; Length 10;  
 Best Local Similarity 62.5%; Pred. No. 27;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 7 DQYGHAAF 14  
 Db 3 DDYGHMRF 10

---

RESULT 5

S47372  
 T-cell antigen receptor VJ junction beta chain - human  
 C;Species: Homo sapiens (man)  
 C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999  
 C;Accession: S47372  
 R;Lechner, P.J.  
 Submitted to the EMBL Data Library, August 1994  
 Submitted to the EMBL Data Library, August 1994  
 A;Cross-references: EMBL:Z35697; NID:g527485; PIDN:CAA84766.1; PID:g527486  
 C;Keywords: T-cell receptor

Query Match 36.0%; Score 31; DB 2; Length 13;  
 Best Local Similarity 71.4%; Pred. No. 36;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RAAYDQY 9  
 Db 6 RSAVEQY 12

---

RESULT 6

S47384  
 T-cell antigen receptor VJ junction beta chain - human  
 C;Species: Homo sapiens (man)  
 C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999  
 C;Accession: S47384  
 R;Lechner, P.J.  
 Submitted to the EMBL Data Library, August 1994  
 Submitted to the EMBL Data Library, August 1994  
 A;Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T ce  
 A;Reference number: S47385  
 A;Accession: S47384  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-13 <LEH>  
 A;Cross-references: EMBL:Z35704; NID:g527501; PIDN:CAA84773.1; PID:g527502  
 C;Keywords: T-cell receptor

Query Match 36.0%; Score 31; DB 2; Length 13;  
 Best Local Similarity 71.4%; Pred. No. 36;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RAAYDQY 9  
 Db 6 RSAVEQY 12

---

RESULT 7

GMROL  
 Leucosulfakinin - Madeira cockroach  
 N;Alternate names: LSK  
 C;Species: Leucophaea maderae (Madeira cockroach)  
 C;Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 13-Sep-1996  
 C;Accession: A01622  
 R;Nachman, R.J.; Holman, G.M.; Haddon, W.F.; Ling, N.  
 Science 234, 71-73, 1986  
 A;Title: Leucosulfakinin, a sulfated insect neuropeptide with homology to gastrin and chd  
 A;Reference number: A01622; MUID:86315858; PMID:3749893  
 A;Accession: A01622  
 A;Molecule type: protein  
 A;Residues: 1-11 <NAC>  
 C;Superfamily: gastrin  
 C;Keywords: amidated carboxyl end; hormone; sulfoprotein  
 F;6/Binding site: sulfate (Tyr) (covalent) #status experimental  
 F;11/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 34.9%; Score 30; DB 1; Length 11;  
 Best Local Similarity 44.4%; Pred. No. 45;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 6 YDQYGHAAF 14  
 Db 3 FDDYGHMRF 11

---

RESULT 8

S47368  
 T-cell antigen receptor VJ junction beta chain - human

Copyright (c) 1993 - 2004 Compugen Ltd.	GenCore version 5.1.6									
OM protein - protein search, using sw model										
Run on:	July 11, 2004, 09:34:09 ; Search time 11 Seconds (without alignments)									
Title:	75.738 Million cell updates/sec									
Perfect score:	US-10-001-938-3									
Sequence:	1 QKRAAYDQYGHAAFFEQ 16									
Scoring table:	BLOSUM62									
Gapop 10.0 , Gapext 0.5										
Searched:	141681 seqs, 52070155 residues									
Total number of hits satisfying chosen parameters:	880									
Minimum DB seq length: 0										
Maximum DB seq length: 16										
Post-processing:	Minimum Match 0% Maximum Match 100% Listing first 45 summaries									
Database :	SwissProt_42;*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.										
SUMMARIES										
Result No.	Score	Query Match	Length	DB ID	Description	RESULT 1	NSKL_SARBU	STANDARD;	PRT;	9 AA.
-	-	-	-	-	-	ID NSKL_SARBU	AC P41492;	DT 01-NOV-1995 (Rel. 32, Created)	DT 01-FEB-1996 (Rel. 33, Last sequence update)	DR Neosulfakinin-I (NEB-SK-I).
1	34	39.5	9	1 NSK1_SARBU	P41492 sarcophaga	ID NSKL_SARBU	AC P36885;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	DE Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).	
2	34	39.5	11	1 LSKP_PERAM	P36885 sarcophaga	ID NSKL_SARBU	AC P41493;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea; Sarcophagidae; Sarcophaga.	
3	39.5	14	1 NSK2_SARBU	PS00259; GASTRIN; 1; PROSIT; IPR001651; Gastrin.	P09039 leucophaea	ID NSKL_SARBU	AC P04428;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	OC NCBI_TaxID=7385;	
4	38.4	12	1 LSK2_LEUMA	NP1_LYMST	P04428 leucophaea	ID NSKL_SARBU	AC P80178;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	OX [1]	
5	36.0	10	1 LSK1_LEUMA	NP2_LYMST	P80178 lymanea sta	ID NSKL_SARBU	AC P80179;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RC TISSUE=Head;	
6	34.9	11	1 LSK1_LEUMA	NP3_LYMST	P80180 lymanea sta	ID NSKL_SARBU	AC P80180;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.; "Isolation and primary structure of two sulfakinin-like peptides from the fleshfly, <i>Neobellieria bullata</i> ."	
7	29.1	13	1 NP1_LYMST	NP4_LYMST	P80181 lymanea sta	ID NSKL_SARBU	AC P80181;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT Comp. Biochem. Physiol. 103C:135-142 (1992).	
8	26.7	13	1 NP2_LYMST	NP5_LYMST	P80182 canis famil	ID NSKL_SARBU	AC P80182;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT TISSUE=Head;	
9	25.6	13	1 NP3_LYMST	NP5_LYMST	P80183 lymanea sta	ID NSKL_SARBU	AC P80183;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	
10	24.4	13	1 PSBP_PINPS	BCDA_LYMDI	P80184 pinus pinas	ID NSKL_SARBU	AC P80184;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.; "Isolation and primary structure of two sulfakinin-like peptides from the fleshfly, <i>Neobellieria bullata</i> ."	
11	24.4	13	1 PSBP_PINPS	NUDM_CANFA	P80185 lymantria d	ID NSKL_SARBU	AC P80185;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	
12	24.4	15	1 PSBP_PINPS	NP5_LYMST	P80186 canis famil	ID NSKL_SARBU	AC P80186;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.; "Isolation and primary structure of two sulfakinin-like peptides from the fleshfly, <i>Neobellieria bullata</i> ."	
13	22.1	12	1 PSBP_PINPS	MY14_EISFO	P80187 lymanea sta	ID NSKL_SARBU	AC P80187;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	
14	22.1	13	1 PSBP_PINPS	NP5_LYMST	P80188 eisenia foie	ID NSKL_SARBU	AC P80188;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.; "Isolation and primary structure of two sulfakinin-like peptides from the fleshfly, <i>Neobellieria bullata</i> ."	
15	22.1	14	1 PSBP_PINPS	UC14_MAIZE	P80189 zea mays (m)	ID NSKL_SARBU	AC P80189;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	
16	22.1	15	1 PSBP_PINPS	ARCD_PSEPU	P80190 pseudomomas	ID NSKL_SARBU	AC P80190;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.; "Isolation and primary structure of two sulfakinin-like peptides from the fleshfly, <i>Neobellieria bullata</i> ."	
17	22.1	16	1 PSBP_PINPS	LSPN05_PINPS	P80191 pinus pinas	ID NSKL_SARBU	AC P80191;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	
18	20.9	7	1 PSP3_PHYP4	RS19_CLYEP	P80192 physcomitre	ID NSKL_SARBU	AC P80192;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.; "Isolation and primary structure of two sulfakinin-like peptides from the fleshfly, <i>Neobellieria bullata</i> ."	
19	20.9	12	1 PSP3_PHYP4	RS19_CLYEP	P80193 clover yell	ID NSKL_SARBU	AC P80193;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	
20	20.9	12	1 PSP3_PHYP4	RS19_CLYEP	P80194 elmy Yellows	ID NSKL_SARBU	AC P80194;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.; "Isolation and primary structure of two sulfakinin-like peptides from the fleshfly, <i>Neobellieria bullata</i> ."	
21	20.9	12	1 PSP3_PHYP4	RS19_CLYEP	P80195 tenebrio mo	ID NSKL_SARBU	AC P80195;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	
22	20.9	14	1 ADFA_TENMO	RS19_CLYEP	P80196 clover prol	ID NSKL_SARBU	AC P80196;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.; "Isolation and primary structure of two sulfakinin-like peptides from the fleshfly, <i>Neobellieria bullata</i> ."	
23	20.9	14	1 ADFA_TENMO	RS19_CLYEP	P80197 locust witc	ID NSKL_SARBU	AC P80197;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	
24	20.9	14	1 ADFA_TENMO	RS19_CLYEP	P80198 pinus pinas	ID NSKL_SARBU	AC P80198;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.; "Isolation and primary structure of two sulfakinin-like peptides from the fleshfly, <i>Neobellieria bullata</i> ."	
25	20.9	15	1 ACT_FINPS	FIBA_ANAPL	P80199 anas platyrhynchos	ID NSKL_SARBU	AC P80199;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	
26	20.9	15	1 ACT_FINPS	FIBA_ANAPL	P80200 zea mays (m)	ID NSKL_SARBU	AC P80200;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.; "Isolation and primary structure of two sulfakinin-like peptides from the fleshfly, <i>Neobellieria bullata</i> ."	
27	20.9	15	1 ACT_FINPS	FIBA_ANAPL	P80201 pseudomonas	ID NSKL_SARBU	AC P80201;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	
28	20.9	16	1 ALRX_PSEPU	MLB_SQUAC	P80202 squalus aca	ID NSKL_SARBU	AC P80202;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.; "Isolation and primary structure of two sulfakinin-like peptides from the fleshfly, <i>Neobellieria bullata</i> ."	
29	20.9	16	1 ALRX_PSEPU	MLB_SQUAC	P16093 trichosanth	ID NSKL_SARBU	AC P16093;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	
30	20.9	16	1 ALRX_PSEPU	MLB_SQUAC	P82240 oncorhynchus	ID NSKL_SARBU	AC P82240;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.; "Isolation and primary structure of two sulfakinin-like peptides from the fleshfly, <i>Neobellieria bullata</i> ."	
31	20.9	17	1 ALRX_PSEPU	MLB_SQUAC	P82470 schistocerc	ID NSKL_SARBU	AC P82470;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	
32	19.8	14	1 TKN1_SACHGR	TKN1_SACHGR	P59683 bacillus st	ID NSKL_SARBU	AC P59683;	DT 01-FEB-1996 (Rel. 33, Last sequence update)	RT MEDLINE=93083101; PubMed=1360367;	

RA Veenstra J.A.;  
 RT "Isolation and structure of two gastrin/CCK-like neuropeptides from  
 RL the American cockroach homologous to the leucosulfakinins.";  
 CC  
 -!- FUNCTION: Stimulates hindgut contractions.  
 DR PIR; A60656; A60656.  
 DR InterPro; IPR001651; Gastrin.  
 DR PROSITE; PS00259; GASTRIN; 1.  
 KW Hormone; Amidation; Sulfation.  
 FT MOD\_RES 6 6 SULFATION.  
 FT MOD\_RES 11 11 AMIDATION.  
 SQ SEQUENCE 11 AA; 1445 MW; 8B4E0680E86B5AAA CRC64;

Query Match 39.5%; Score 34; DB 1; Length 11;  
 Best Local Similarity 55.6%; Pred. No. 5;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 YDQYGHAAF 14

Db 3 FDDYGHMRF 11

RESULT 3

ID NSK2\_SARBU STANDARD; PRT; 14 AA.

AC P41493;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-FEB-1995 (Rel. 33, Last annotation update)  
 DE Neosulfakinin-II (NEB-SK-II).  
 OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
 OC Sarcophagidae; Sarcophaga.  
 RN NCBI\_TaxID=7385;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Head;  
 RX MEDLINE=93083101; PubMed=1360367;  
 RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.;  
 RT the fleshfly, *Neobellieria bullata*.;  
 RT Comp. Biochem. Physiol. 103C:135-142(1992).  
 CC -!- FUNCTION: Myotrophic peptide.  
 CC -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.  
 DR InterPro; IPR001651; Gastrin.  
 DR PROSITE; PS00259; GASTRIN; 1.  
 KW Neuropeptide; Amidation; Sulfation.  
 FT MOD\_RES 9 9 SULFATION (POTENTIAL).  
 FT MOD\_RES 14 14 AMIDATION (POTENTIAL).  
 SQ SEQUENCE 14 AA; 1794 MW; 8B4E06D5B61C62AA CRC64;

Query Match 39.5%; Score 34; DB 1; Length 14;  
 Best Local Similarity 55.6%; Pred. No. 6.5;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 YDQYGHAAF 14

Db 6 FDDYGHMRF 14

RESULT 4

ID LOSK\_LOCMT STANDARD; PRT; 12 AA.

AC P47733;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Sulafakinin (LOM-SK).  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acrideromorpha;  
 OC Acriidoidea; Acriidae; Oedipodinae; Locusta.  
 OC NCBI\_TaxID=7004;  
 RN [1]

RP SEQUENCE.

RC TISSUE=Brain;  
 RA Schoofs L., Holman G.L., Hayes T.K., Nachman R.J., de Loof A.;  
 RL (In) McCaffery A., Wilson I. (eds.); Chromatography and isolation of insect hormones and pheromones,  
 RL pp.231-241, Plenum Press, New York (1990).  
 CC -!- FUNCTION: Myotropic peptide.  
 CC -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.  
 DR InterPro; IPR001651; Gastrin.  
 DR PROSITE; PS00259; GASTRIN; 1.  
 KW Hormone; Amidation; Sulfation; Pyrrolidone carboxylic acid.  
 FT MOD\_RES 1 1 SULFATION (POTENTIAL).  
 FT MOD\_RES 7 7 AMIDATION.  
 SQ SEQUENCE 12 AA; 1440 MW; 9B5B5DA9BD6B5AAA CRC64;

Query Match 38.4%; Score 33; DB 1; Length 12;  
 Best Local Similarity 60.0%; Pred. No. 8.2;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 YDQYGHAAF 14

Db 3 ASDDYGHMRF 12

RESULT 5

ID LSK2\_LEUMA STANDARD; PRT; 10 AA.

AC P09039;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Leucosulfakinin-II (LSK-II).  
 OS Leucophaea maderae (Madeira cockroach), and  
 Periplaneta americana (American cockroach), and  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
 OC Blaberidae; Leucophaea.  
 RN NCBI\_TaxID=6988, 6978;  
 RN [1]  
 RP SEQUENCE.

RC SPECIES=L.maderae;  
 RX MEDLINE=87048769; PubMed=3778455;  
 RA Nachman R.J.; Holman G.M.; Cook B.J.; Haddon W.F.; Ling N.;  
 RT "Leucosulfakinin-II, a blocked sulfated insect neuropeptide with  
 RL homology to cholecystokinin and gastrin";  
 RL Biochem. Biophys. Res. Commun. 140:357-364(1986).  
 RN [2]  
 RP SEQUENCE.  
 RC SPECIES=P.americana; TISSUE=Corpora cardiaca;  
 RX MEDLINE=90137190; PubMed=2615921;  
 RA "Isolation and structure of two gastrin/CCK-like neuropeptides from  
 RT the American cockroach homologous to the leucosulfakinins.";  
 RL Neuropeptides 14:145-149(1989).  
 CC -!- FUNCTION: CHANGE THE FREQUENCY AND AMPLITUDE OF CONTRACTIONS OF  
 CC THE COCKROACH HINGUT. STIMULATES MUSCLE CONTRACTION OF HINDGUT.  
 CC -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.  
 DR PIR; A26335; GMROL2.  
 DR InterPro; IPR001651; Gastrin.  
 DR PROSITE; PS00259; GASTRIN; 1.  
 DR Hormone; Amidation; Sulfation; Pyrrolidone carboxylic acid.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 5 5 SULFATION (IN L.MADERAE, BUT NOT IN  
 P.AMERICA).  
 FT MOD\_RES 10 10 AMIDATION.  
 SQ SEQUENCE 10 AA; 1255 MW; 9B4F5391E86B5AAA CRC64;

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:33:34 ; Search time 52 Seconds  
 (without alignments)  
 86.938 Million cell updates/sec

Title: US-10-001-938-3

Perfect score: 86

Sequence: 1\_QKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62

Gapext 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 459173

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing First 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp1980s:\*

2: geneseqp1990s:\*

3: geneseqp2000s:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

Result No.	Score	Query %	Match Length	DB ID	Description
1	86	100.0	16	5 AAU98852	Aau98852 E.Coli DN
2	86	100.0	16	6 ABR55126	Abr55126 E. coli d
3	81	94.2	15	2 AAW25795	AAW25795 Peptide d
4	81	94.2	15	3 AAY79543	Aay79543 E. coli d
5	81	94.2	15	4 AAM99341	Aam99341 Vaccine r
6	81	94.2	15	5 AAU09842	Aau09842 E.coli dn
7	81	94.2	15	5 AAU98876	Aau98876 E. coli D
8	81	94.2	15	5 AAE19457	Aae19457 Human hea
9	81	94.2	15	6 ABR55132	Abr55132 E. coli D
10	81	94.2	16	4 AAM99344	Aam99344 Vaccine r
11	72	83.7	15	2 AAR95446	Aar95446 RA suscep
12	72	83.7	15	2 AAW25796	AAW25796 Peptide d
13	72	83.7	15	5 AAE19458	Aae19458 Heat shoc
14	72	83.7	15	6 ABR55133	Abr55133 E. coli d
15	64.5	75.0	14	2 AAR95445	Aar95445 RA suscep
16	54	62.8	15	2 AAW25798	AAW25798 Peptide d
17	54	62.8	15	6 ABR55123	Abr55123 E. coli d
18	34	39.5	9	4 AAU03347	Aau03347 Fruit fly
19	34	39.5	9	4 AAU03346	Aau03346 Fruit fly
20	34	39.5	9	4 AAU03351	Aau03351 Fruit fly
21	34	39.5	9	7 ADE14680	Adel14680 Drosulfak
22	34	39.5	9	7 ADE14676	Adel14676 DmGPCR bi
23	34	39.5	9	7 ADE14679	Adel14679 DmGPCR bi
24	34	39.5	9	7 ADE14677	Adel14677 DmGPCR bi
25	34	39.5	14	7 AAB14081	Aab14081 Human HLA

**ALIGNMENTS**

RESULT 1

ID AAU98852 standard; peptide; 16 AA.

XX AAU98852;

XX DT 22-AUG-2002 (First entry)

XX DE E.Coli DNAJ 61 immunogenic peptide.

XX KW Immunogenic peptide; heat shock protein; HSP; DNAJ; immunomodulatory; autoimmune disease; arthritidis; articular juvenile idiopathic arthritis; cytostatic; antiinflammatory; antibacterial; antiarthritic; KW infectious disease; inflammatory bowel disease; cancer; mucosal tolerisation; DNA vaccination; anergy induction.

XX OS Escherichia coli.

XX PN WO200236611-A2.

XX PD 10-MAY-2002.

XX PP 31-OCT-2001; 2001WO-US045344.

XX PR 01-NOV-2000; 2000US-0245181P.

(REGC ) UNIV CALIFORNIA.  
 (MART/) MARTINI A.

XX PA PI Martini A, Albani S, Carson DA, Prakken BJ;

XX DR WPI; 2002-489999/52.

PT New immunomodulatory peptides from heat shock proteins, useful for treating immunological disorder in subjects such as humans, e.g. autoimmune disease (e.g. arthritis), infectious disease, inflammatory bowel disease or cancer.

PT autoimmunity, cytostatic, antiinflammatory, antibacterial or immunomodulatory, cytostatic, antiinflammatory, antibacterial or antiarthritic properties and can stimulate expression of interleukins, tumour necrosis factor and transforming growth factor beta. The immunogenic peptide is useful for modulating (i.e. augmenting/inducing or reducing/inhibiting) an immune response in a subject having an

26 34 39.5 14 4 AAU03353 Aau03353 Fruit fly

27 34 39.5 14 6 AAG79793 Aag79793 Peptide 6

28 34 39.5 14 7 ADE14681 Adel14681 Drosulfak

29 34 39.5 15 2 AAR10086 Aar10086 HLA Dw4 i

30 34 39.5 15 2 AAR10089 Aar10089 HLA Dw4 p

31 34 39.5 15 2 AAW25799 Aaw25799 Synthetic

32 34 39.5 15 2 AAW25800 Abr55139 Human S1

33 34 39.5 15 3 AAY79544 Aay79544 Human leu

34 34 39.5 15 4 AAM99340 Aam99340 Vaccine r

35 34 39.5 15 5 AAU09843 Aau09843 Human leu

36 34 39.5 15 6 ABR55138 Abr55138 Human S1

37 34 39.5 15 6 ABR55139 Abr55139 Human S2

38 34 39.5 15 6 AAG79792 Aag79792 Peptide 6

39 31 36.0 12 3 AAB39771 Ab39771 Anti-IL12

40 31 36.0 12 3 AAB39769 Ab39769 Anti-IL12

41 31 36.0 12 3 AAB39510 Ab39510 Anti-HL-

42 31 36.0 14 3 AAB13800 Aab13800 Human HLA

43 31 36.0 14 3 AAB14089 Aab14089 Human HLA

44 31 36.0 14 3 AAG79798 Aag79798 Peptide 6

CC immunological disorder (e.g. autoimmune disease such as arthritis or  
 CC articular juvenile idiopathic arthritis), an infectious disease, an  
 CC inflammatory bowel disease or cancer. The immunogenic peptide of the  
 CC invention is also useful for modulating immunoefector cell  
 CC responsiveness in a subject. The immunogenic peptide is particularly  
 useful for treating the above-mentioned diseases in mammals, e.g. cat,  
 CC dog, horse, farm animal (e.g. ovine, bovine or porcine) or human. In  
 general, the peptide is useful in methods involving mucosal tolerisation,  
 CC DNA vaccination, energy induction or active immunisation. The present  
 CC sequence represents an E. coli DNAJ immunogenic peptide of the invention  
 CC SQ Sequence 16 AA;

Query Match 100.0%; Score 86; DB 5; Length 16;

Best Local Similarity 100.0%; Pred. No. 4.1e-08; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OKRAAYDQYGHAAFEQ 16  
 Db 1 OKRAAYDQYGHAAFEQ 16

CC invention  
 CC SQ Sequence 16 AA;  
 Query Match 100.0%; Score 86; DB 6; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-08; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 OKRAAYDQYGHAAFEQ 16  
 Db 1 OKRAAYDQYGHAAFEQ 16

RESULT 2  
 ABR55126  
 ID ABR55126 standard; peptide; 16 AA.

XX ABR55126;  
 AC XX  
 DT 03-JUL-2003 (first entry)  
 DE XX  
 DE. coli dnaJ61 antigen-specific epitope peptide.  
 XX  
 KW Antigen-specific epitope; immune response; T cell; cytokine;  
 KW antiarthritic; antidiabetic; neuroprotective; anti-inflammatory;  
 KW cytostatic; antithyroid; anti-asthmatic; immunosuppressive;  
 KW antipsoriatic; anti-ulcer; antianæmic; cardiant; respiratory;  
 KW anti-allergic; dermatological; antipsoriatic.  
 XX  
 OS Escherichia coli.  
 XX  
 PN WO9734002-A1.  
 XX  
 PD 18-SEP-1997.  
 XX  
 PP 20-FEB-1997; 97WO-US002957.  
 XX  
 PR 15-MAR-1996; 96US-00618464.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Carson DA, Albani S;  
 XX  
 DR WPI; 1997-470882/43.  
 XX  
 PT Vaccine for protecting against arthritogenic peptide(s) containing dnaJpl  
 PT - also determining pre-disposition to rheumatoid arthritis by detecting  
 PT anti-arthritogenic peptide antibodies.  
 XX  
 PS Disclosure; Page 5; 44pp; English.

Peptides AAW25795-96 and AAW25798 are derived from the dnaJ protein from Escherichia coli. One peptide, dnaJpl (AAW25795) was found to induce relatively strong immune response in seropositive adult rheumatoid arthritis patients. This peptide was used in a vaccine for protecting against arthritogenic peptides. The vaccine contains a carrier, pure dnaJpl peptide or a recombinant gene expression vector encoding the dnaJpl peptide. The vaccines are used to prevent rheumatoid arthritis (or other autoimmune diseases). Vaccines can target the arthritogenic peptides before they are presented to the systemic immune system. Anti-dnaJpl antibodies can also be used for passive immunisation. Antibodies can be used to detect rheumatoid arthritis-susceptibility sequences on HLA or in populations of E. coli in the gastrointestinal tract

CC The invention relates to a novel method for modulating an immune response  
 CC in a subject having an immune-related disorder. The method comprises: (a)  
 CC administering an antigen-specific epitope, where administration provides  
 CC epitope-specific T cell immune modulation; and (b) administering a  
 CC cytokine, an agent that effects cytokine activity or expression, or an  
 CC antibody therapy. The method of the invention has antiarthritic,  
 CC anticytokine, neuroprotective, anti-inflammatory, cytostatic,  
 CC antithyroid, anti-asthmatic, immunosuppressive, antipsoriatic, anti-  
 CC ulcer, antianaemic, cardiant, respiratory general, antiallergic,  
 CC dermatological, and antipsoriatic activity. The method is useful for  
 CC modulating an immune response in a subject having an immune-related  
 CC disorder. The present sequence is used in the exemplification of the

Query Match 94.2%; Score 81; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-07;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 OKRAAYDQYGHAAFE 15  
 Db 1 OKRAAYDQYGHAAFE 15

GenCore version 5.1.6  
copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:37:19 ; Search time 38 seconds

(without alignments)  
132.850 Million cell updates/sec

Title: US-10-001-938-3  
perfect score: 86  
Sequence: 1 QKRAAYDQYGHAAFEQ 16  
  
Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5  
  
Searched: 1017041 seqs, 315518202 residues  
  
Total number of hits satisfying chosen parameters: 4499  
  
Minimum DB seq length: 0  
Maximum DB seq length: 16  
  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25:  
1: sp\_archea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_mhc:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rabbit:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:  
15: sp\_rvirus:  
16: sp\_bacteriap:  
17: sp\_archeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

17 20 23.3 12 11 Q9QVF2  
18 20 23.3 14 10 Q9FYTO  
19 20 23.3 15 2 Q9JP39  
20 20 23.3 15 10 Q9S8I1  
21 20 23.3 16 10 Q82403  
22 22.7 11 2 P96319  
19.5 22.7 8 6 Q9BF82  
23 19 22.1 8 6 Q9BFC2  
24 19 22.1 8 6 Q9BF90  
25 19 22.1 8 6 Q9BFB1  
26 19 22.1 8 6 Q9BFA1  
27 19 22.1 8 6 Q9BFB87  
28 19 22.1 8 6 Q9BFB9  
29 19 22.1 8 6 Q9BFB5  
30 19 22.1 8 6 Q9BFB8  
31 19 22.1 8 6 Q9BFA8  
32 19 22.1 8 6 Q9BFB3  
33 19 22.1 8 6 Q9BFB2  
34 19 22.1 8 6 Q9BFB5  
35 19 22.1 8 6 Q9BFB6  
36 19 22.1 8 6 Q9BFA9  
37 19 22.1 8 6 Q9BFB8  
38 19 22.1 8 6 Q9BFA9  
39 19 22.1 8 6 Q9BFB4  
40 19 22.1 8 6 Q9BFC3  
41 19 22.1 8 6 Q9BFA4  
42 19 22.1 8 6 Q9BFA2  
43 19 22.1 8 6 Q9BFB5  
44 19 22.1 8 6 Q9BFC1  
45 19 22.1 8 6 Q9BFB9  
8 6 Q9BFB89  
Q9qvf2 rattus sp.  
Q9fyto allium cepa  
Q9jp39 pseudomonas  
Q9s8i1 volvox cart  
P96319 desulfobiv  
Q9bf82 ursus arcto  
Q9bfc2 macropus eu  
Q9bf90 tragelaphus  
Q9bfb1 echinops te  
Q9bfa1 atelles fusc  
Q9bfb8 tapirus ind  
Q9bfb9 euphractus  
Q9bfb8 chaetophrac  
Q9bfa8 loxodonta a  
Q9bfa9 procavia ca  
Q9bfb2 sorex arane  
Q9bfb5 erinaceus c  
Q9bfb6 myrmecophag  
Q9bfb3 condylura c  
Q9bfb8 equus cabal  
Q9bfa9 rousettus l  
Q9bfb4 panthera on  
Q9bfc3 didelphis m  
Q9bfa4 tupaia mino  
Q9bfa2 tarsius ban  
Q9bfb5 leopards p  
Q9bfc1 choloepus h  
Q9bf89 okapia john

#### ALIGNMENTS

RESULT 1  
ID 082407 PRELIMINARY; PRT; 16 AA.  
ID 082407; AC 082407;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE Alcohol dehydrogenase (EC 1.1.1.1) (Fragment).  
GN ADH.  
OS Fragaria viridis (Wild strawberry).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Rosales; Rosaceae; Rosoideae; Fragaria.  
OX NCBI\_TAXID=64942;  
[1]  
RN SEQUENCE FROM N.A.

RC STRAIN=FRA 341;  
RA Yu H., Davis T.M.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF000220; AAC36547; 1; -  
DR GO; GO:0004025; F:alcohol dehydrogenase activity, iron-dependent; IEA.  
DR GO; GO:0004023; F:alcohol dehydrogenase activity, metal ion-<sup>i</sup>-; .; IEA.  
DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
KW Oxidoreductase.  
FT NON\_TER 1 1  
FT SEQUENCE 16 AA; 1682 MW; CERF73B5B28A4D817 CRC64;  
Query Match 27.9%; Score 24; DB 10; Length 16;  
Best Local Similarity 80.0%; Pred. No. 2.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 YGHAA 13  
Db 10 YGHEA 14

RESULT 2

17 20 23.3 12 11 Q9QVF2  
18 20 23.3 14 10 Q9FYTO  
19 20 23.3 15 2 Q9JP39  
20 20 23.3 15 10 Q9S8I1  
21 20 23.3 16 10 Q82403  
22 22.7 11 2 P96319  
19.5 22.7 8 6 Q9BF82  
23 19 22.1 8 6 Q9BFC2  
24 19 22.1 8 6 Q9BF90  
25 19 22.1 8 6 Q9BFB1  
26 19 22.1 8 6 Q9BFA1  
27 19 22.1 8 6 Q9BFB87  
28 19 22.1 8 6 Q9BFB9  
29 19 22.1 8 6 Q9BFB5  
30 19 22.1 8 6 Q9BFB8  
31 19 22.1 8 6 Q9BFA8  
32 19 22.1 8 6 Q9BFB3  
33 19 22.1 8 6 Q9BFB2  
34 19 22.1 8 6 Q9BFB5  
35 19 22.1 8 6 Q9BFB6  
36 19 22.1 8 6 Q9BFA9  
37 19 22.1 8 6 Q9BFB8  
38 19 22.1 8 6 Q9BFA9  
39 19 22.1 8 6 Q9BFB4  
40 19 22.1 8 6 Q9BFC3  
41 19 22.1 8 6 Q9BFA4  
42 19 22.1 8 6 Q9BFA2  
43 19 22.1 8 6 Q9BFB5  
44 19 22.1 8 6 Q9BFC1  
45 19 22.1 8 6 Q9BFB9  
8 6 Q9BFB89  
Q9qvf2 rattus sp.  
Q9fyto allium cepa  
Q9jp39 pseudomonas  
Q9s8i1 volvox cart  
P96319 desulfobiv  
Q9bf82 ursus arcto  
Q9bfc2 macropus eu  
Q9bf90 tragelaphus  
Q9bfb1 echinops te  
Q9bfa1 atelles fusc  
Q9bfb8 tapirus ind  
Q9bfb9 euphractus  
Q9bfb8 chaetophrac  
Q9bfa8 loxodonta a  
Q9bfa9 procavia ca  
Q9bfb2 sorex arane  
Q9bfb5 erinaceus c  
Q9bfb6 myrmecophag  
Q9bfb3 condylura c  
Q9bfb8 equus cabal  
Q9bfa9 rousettus l  
Q9bfb4 panthera on  
Q9bfc3 didelphis m  
Q9bfa4 tupaia mino  
Q9bfa2 tarsius ban  
Q9bfb5 leopards p  
Q9bfc1 choloepus h  
Q9bf89 okapia john

17 20 23.3 12 11 Q9QVF2  
18 20 23.3 14 10 Q9FYTO  
19 20 23.3 15 2 Q9JP39  
20 20 23.3 15 10 Q9S8I1  
21 20 23.3 16 10 Q82403  
22 22.7 11 2 P96319  
19.5 22.7 8 6 Q9BF82  
23 19 22.1 8 6 Q9BFC2  
24 19 22.1 8 6 Q9BF90  
25 19 22.1 8 6 Q9BFB1  
26 19 22.1 8 6 Q9BFA1  
27 19 22.1 8 6 Q9BFB87  
28 19 22.1 8 6 Q9BFB9  
29 19 22.1 8 6 Q9BFB5  
30 19 22.1 8 6 Q9BFB8  
31 19 22.1 8 6 Q9BFA8  
32 19 22.1 8 6 Q9BFB3  
33 19 22.1 8 6 Q9BFB2  
34 19 22.1 8 6 Q9BFB5  
35 19 22.1 8 6 Q9BFB6  
36 19 22.1 8 6 Q9BFA9  
37 19 22.1 8 6 Q9BFB8  
38 19 22.1 8 6 Q9BFA9  
39 19 22.1 8 6 Q9BFB4  
40 19 22.1 8 6 Q9BFC3  
41 19 22.1 8 6 Q9BFA4  
42 19 22.1 8 6 Q9BFA2  
43 19 22.1 8 6 Q9BFB5  
44 19 22.1 8 6 Q9BFC1  
45 19 22.1 8 6 Q9BFB9  
8 6 Q9BFB89  
Q9qvf2 rattus sp.  
Q9fyto allium cepa  
Q9jp39 pseudomonas  
Q9s8i1 volvox cart  
P96319 desulfobiv  
Q9bf82 ursus arcto  
Q9bfc2 macropus eu  
Q9bf90 tragelaphus  
Q9bfb1 echinops te  
Q9bfa1 atelles fusc  
Q9bfb8 tapirus ind  
Q9bfb9 euphractus  
Q9bfb8 chaetophrac  
Q9bfa8 loxodonta a  
Q9bfa9 procavia ca  
Q9bfb2 sorex arane  
Q9bfb5 erinaceus c  
Q9bfb6 myrmecophag  
Q9bfb3 condylura c  
Q9bfb8 equus cabal  
Q9bfa9 rousettus l  
Q9bfb4 panthera on  
Q9bfc3 didelphis m  
Q9bfa4 tupaia mino  
Q9bfa2 tarsius ban  
Q9bfb5 leopards p  
Q9bfc1 choloepus h  
Q9bf89 okapia john

O82404 PRELIMINARY; PRT; 16 AA.  
 ID O82404;  
 AC 082404;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Alcohol dehydrogenase (EC 1.1.1.1) (Fragment).  
 GN ADH.  
 OS Fragaria nilgerrensis.  
 OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euerosids I; Rosales; Rosaceae; Rosoideae; Fragaria.  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=Berlin 1;  
 RA Yu H.; Davis T.M.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF000217; AAC36544.1; -.  
 DR GO; GO:0004025; F:alcohol dehydrogenase activity, iron-dependent; IEA.  
 DR GO; GO:0004023; F:alcohol dehydrogenase activity, zinc-dependent; IEA.  
 DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 KW Oxidoreductase.  
 FT NON-TER 1  
 SQ 16 AA; 16 MW; CEF73B5B28A4D952 CRC64;

Query Match 27.9%; Score 24; DB 10; Length 16;  
 Best Local Similarity 80.0%; Pred. No. 2.9e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 YGHAA 13  
 |||||  
 Db 10 YGHEA 14

RESULT 3  
 O82405 PRELIMINARY; PRT; 16 AA.  
 ID O82406;  
 AC 082406;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Alcohol dehydrogenase (EC 1.1.1.1) (Fragment).  
 GN ADH.  
 OS Fragaria moschata.  
 OC Spermatophyta; Streptophyta; Embryophyta; Tracheophyta;  
 OC euerosids I; Rosales; Rosaceae; Rosoideae; Fragaria.  
 OC NCBI\_TaxID=64940;  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=FRA 157;  
 RA Yu H.; Davis T.M.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF000219; AAC36546.1; -.  
 DR GO; GO:0004025; F:alcohol dehydrogenase activity, iron-dependent; IEA.  
 DR GO; GO:0004023; F:alcohol dehydrogenase activity, metal ion-1.; IEA.  
 DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 KW Oxidoreductase.  
 FT NON-TER 1  
 SQ 16 AA; 16 MW; CEF73B5B28A4D952 CRC64;

Query Match 27.9%; Score 24; DB 10; Length 16;  
 Best Local Similarity 80.0%; Pred. No. 2.9e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 YGHAA 13  
 |||||  
 Db 10 YGHEA 14

RESULT 4  
 O82402 PRELIMINARY; PRT; 16 AA.  
 ID O82402;  
 AC 082402;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Alcohol dehydrogenase (EC 1.1.1.1) (Fragment).  
 GN ADH.  
 OS Fragaria nubicola.  
 OC Spermatophyta; Magnoliophyta; eudicots; rosids; euerosids I; Rosales; Rosaceae; Rosoideae; Fragaria.  
 OC NCBI\_TaxID=60188;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=FRA520;  
 RA Yu H.; Davis T.M.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF000213; AAC36540.1; -.  
 DR GO; GO:0004025; F:alcohol dehydrogenase activity, iron-dependent; IEA.  
 DR GO; GO:0004023; F:alcohol dehydrogenase activity, zinc-dependent; IEA.  
 DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 KW Oxidoreductase.  
 FT NON-TER 1  
 SQ 16 AA; 16 MW; CEF73B5B28A4D817 CRC64;

Query Match 27.9%; Score 24; DB 10; Length 16;  
 Best Local Similarity 80.0%; Pred. No. 2.9e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 YGHAA 13  
 |||||  
 Db 10 YGHEA 14

RESULT 5  
 O82781 PRELIMINARY; PRT; 16 AA.  
 ID O82781;  
 AC 082781;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Alcohol dehydrogenase (EC 1.1.1.1) (Aldehyde reductase) (Fragment).  
 GN ADH.  
 OS Fragaria vesca (Woodland strawberry).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicots; rosids; euerosids I; Rosales; Rosaceae; Rosoideae; Fragaria.  
 OC NCBI\_TaxID=57918;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=YELLOW WONDER, and BARON SOLEMACHER;  
 RA Yu H.; Davis T.M.;  
 RT "Genetic relationships among Fragaria species based on RAPDs and an alcohol dehydrogenase (ADH) gene.";  
 RL Genome 0:0-0(1997).  
 CC -!- CATALYTIC ACTIVITY: AN ALCOHOL + NAD(+) = AN ALDEHYDE OR KETONE + NADH.  
 CC -!- COFACTOR: ZINC OR IRON.  
 DR EMBL; AF000216; AAC36543.1; -.  
 DR EMBL; AF000214; AAC36541.1; -.  
 DR GO; GO:0004025; F:alcohol dehydrogenase activity, iron-dependent; IEA.  
 DR GO; GO:0004023; F:alcohol dehydrogenase activity, metal ion-1.; IEA.  
 DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 KW Oxidoreductase.  
 FT NON-TER 1  
 FT NON-TER 16 16

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on:

July 11, 2004, 09:37:59 ;

Search time 23 seconds  
 (without alignments)

35.914 Million cell updates/sec

Title: US-10-001-938-3  
 Perfect score: 86  
 Sequence: 1 QKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 152138

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:  
 \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	81	94.2	15	1 US-08-618-464-4	Sequence 4, Appli
2	81	94.2	15	3 US-09-107-615-4	Sequence 4, Appli
3	81	94.2	15	5 PCT-US95-04896-4	Sequence 4, Appli
4	72	83.7	15	1 US-08-618-464-5	Sequence 5, Appli
5	72	83.7	15	3 US-09-107-615-5	Sequence 5, Appli
6	72	83.7	15	5 PCT-US95-04896-5	Sequence 5, Appli
7	54	62.8	15	1 US-08-618-464-6	Sequence 6, Appli
8	54	62.8	15	3 US-09-107-615-6	Sequence 6, Appli
9	34	39.5	15	1 US-08-618-464-7	Sequence 7, Appli
10	34	39.5	15	1 US-08-618-464-8	Sequence 8, Appli
11	34	39.5	15	3 US-09-107-615-7	Sequence 7, Appli
12	34	39.5	15	3 US-09-107-615-8	Sequence 8, Appli
13	31	36.0	12	1 US-07-732-114A-5	Sequence 5, Appli
14	31	36.0	12	1 US-08-170-114A-5	Sequence 5, Appli
15	31	36.0	14	2 US-08-323-686-3	Sequence 3, Appli
16	31	36.0	14	2 US-08-323-686-26	Sequence 26, Appli
17	31	36.0	15	4 US-09-255-501-67	Sequence 67, Appli
18	31	36.0	15	4 US-09-255-501-68	Sequence 68, Appli
19	30	34.9	9	4 US-09-721-870-180	Sequence 180, APP
20	30	34.9	15	3 US-08-467-023-255	Sequence 255, APP
21	28	32.6	11	3 US-08-467-580-14	Sequence 14, Appli
22	28	32.6	11	5 PCT-US95-08516-14	Sequence 14, Appli
23	28	32.6	15	2 US-08-521-871A-3	Sequence 3, Appli
24	27	31.4	12	1 US-07-732-114A-11	Sequence 11, Appli
25	27	31.4	12	1 US-08-170-114A-11	Sequence 11, Appli
26	27	31.4	13	1 US-08-487-568-45	Sequence 45, Appli
27	27	30.2	6	1 US-07-995-503A-12	Sequence 12, Appli

## ALIGNMENTS

RESULT 1	US-08-618-464-4	SEQUENCE CHARACTERISTICS:
		LENGTH: 15 amino acids
		TYPE: amino acid
		STRANDEDNESS: single
		TOPOLOGY: linear
		MOLECULE TYPE: peptide
		IMMEDIATE SOURCE:
		CLONE: Immunogenic dnaJ Peptide
		FEATURE:
		NAME/KEY: Peptide
		LOCATION: 1..15
Query Match	94.2%; Score 81; DB 1; Length 15;	

Best Local Similarity 100.0%; Pred. No. 5.8e-08; Mismatches 0; Indels 0; Gaps 0;

Matches 15; Conservative 0;

QY 1 QKRAAYDQYGHAAFE 15

Db 1 QKRAAYDQYGHAAFE 15

RESULT 2  
US-09-107-615-4  
Sequence 4, Application US/09107615; Patent No. 6153200  
GENERAL INFORMATION:

; APPLICANT: CARSON, DENNIS A.

; APPLICANT: ALBANI, SALVATORE

; TITLE OF INVENTION: VACCINE COMPOSITIONS AND METHODS USEFUL IN

; TITLE OF INVENTION: INDUCING IMMUNE PROTECTION AGAINST

; TITLE OF INVENTION: ARTHRITIC PEPTIDES INVOLVED IN

; TITLE OF INVENTION: THE PATHOGENESIS OF RHEUMATOID ARTHRITIS

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FISH &amp; RICHARDSON P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: CALIFORNIA

; COUNTRY: US

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/04896

; APPLICATION NUMBER: US/09/107,615

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: HOWELLS, STACY L.

; REFERENCE NUMBER: 34,842

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619/678-5070

; TELEFAX: 619/678-5099

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; NAME/KEY: Peptide

; LENGTH: 15 amino acids

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: Peptide

; IMMEDIATE SOURCE:

; CLONE: Immunogenic dnaJ Peptide

; FEATURE:

; NAME/KEY: Peptide

; LOCATION: 1..15

; PCT-US95-04896-4

; RESULT 4  
US-08-618-464-5

; Sequence 5, Application US/08618464

; Patent No. 5773570

; GENERAL INFORMATION:

; APPLICANT: CARSON, DENNIS A.

; APPLICANT: ALBANI, SALVATORE

; TITLE OF INVENTION: VACCINE COMPOSITIONS AND METHODS USEFUL IN

; TITLE OF INVENTION: INDUCING IMMUNE PROTECTION AGAINST ARTHRITIC

; TITLE OF INVENTION: INVOLVED IN THE PATHOGENESIS OF RHEUMATOID ARTHRITIS

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FISH &amp; RICHARDSON P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: CALIFORNIA

; COUNTRY: US

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; GENERAL INFORMATION:

; Sequence 4, Application PC/TUS9504896

; GENERAL INFORMATION: